## REMARKS/ARGUMENTS

In this reply, Claims 23, 37, 38, and 41-50 are amended to be presented in a condition for allowance as discussed in the interview this morning. Computer-readable storage medium Claims 41, 47, and 48 have been converted to independent claims that reflect the amendments to method Claims 23, 37, and 38, respectively. Claims 42-46 have been amended to depend from converted Claim 41, and Claims 49-50 have been amended to depend from converted Claim 47. Claims 23-24, 26-29, and 37-50 are pending in the application. Applicant respectfully requests that the amendments be entered to efficiently advance this application towards a Notice of Allowance, as discussed in the interview.

### INTERVIEW

Examiner Sangwoo Ahn is thanked for his thoughtful participation in the interview this morning. The proposed and discussed amendments are reflected in Claims 23, 37, and 38. Claims 23 and 38 have been recognized as containing allowable subject matter. Accordingly, this reply focuses on the differences between Claim 37 and the art of record. As requested, Applicants provide a detailed explanation herein of the fundamental difference between a selection of an occurrence counting technique based on "how busy a computer system . . . currently is," "before the frequent itemset operation is performed," as recited in Claim 37, and the determinations based on "data characteristics" (Agrawal1, col. 11, ln. 39) and "data structures" (Agrawal2, col. 2, ln. 61) that are mentioned in the references.

### INDEPENDENT CLAIM 23

The Examiner agrees with Applicants that Claim 23 contains allowable subject matter. As discussed in the interview, the amendments to Claim 23 do not affect the reasons for

allowability. Applicant respectfully requests that the amendments be entered to efficiently advance this application towards a Notice of Allowance.

Claim 23 features a method for "dynamically selecting" an occurrence counting technique based on cost estimates of available occurrence counting techniques. For a prefix tree technique, generating a cost estimate comprises: "determining a size of a candidate prefix tree; determining an amount of memory that can be used for the candidate prefix tree; comparing the size of the candidate prefix tree to the amount of memory that can be used to store the candidate prefix tree; and generating an I/O cost estimate for the prefix tree technique based, at least in part, on the size of the candidate prefix tree and the amount of memory that can be used to store the candidate prefix tree." Claim 23 also recites: "selecting the occurrence counting technique that has the lowest cost estimate." The art of record does not show at least the recited features, and Claim 23 is in condition for allowance.

CLAIMS 24, 26-29, AND 41-46

Claims 24 and 26-29 depend from Claim 23, and the Examiner agrees that Claims 24 and 26-29 contain allowable subject matter for similar reasons as those discussed with respect to Claim 23. Claim 41-46 are computer-readable storage medium claims that include the allowable subject matter recited in Claim 23. Thus, Claims 41-46 are allowable over the art of record for at least those reasons provided with respect Claim 23.

INDEPENDENT CLAIM 37

Claim 37 was rejected under 35 U.S.C. § 103(a) as being allegedly unpatentable over U.S. Patent No. 6,324,533 ("Agrawal") in view of U.S. Patent No. 6,513,029 ("Agrawal2"). This rejection is respectfully traversed.

Claim 37 recites a method for "dynamically selecting" an occurrence counting technique "based on conditions existing before a frequent itemset operation is performed in a computing environment in which the frequent itemset operation is to be performed." The "conditions include how busy a computer system in which the frequent itemset operation is to be performed currently is." The selected occurrence counting technique is used "to count occurrences of at least one combination to determine whether said at least one combination satisfies frequency criteria" that is associated with a frequent itemset operation.

At best, Agrawal1 features a "choice of the best SQL-OR approach" that "depends on a number of data characteristics like the number of items, total number of transactions, average length of a transaction, etc." (col. 11, ln. 37-40). **The choice in Agrawal1 depends on data characteristics, "as defined in Table 1" (col. 11, ln. 44), not on "conditions existing before a frequent itemset operation is performed in a computing environment in which the frequent itemset operation is to be performed," as recited in Claim 37**. Further, **the choice in Agrawal1 has nothing to do with "how busy a computer system in which the frequent itemset operation is to be performed currently is," before the frequent itemset operation is performed**. Finally, **Agrawal1 does not involve "dynamically selecting" an occurrence counting technique at all. Instead, Agrawal1 mentions a technique for analyzing test results to find the best static approach to use (col. 14, ln. 36-46)**.

**The data characteristics of Agrawal1 are not how busy a computer system currently is.** As defined in Table 1, the data characteristics of Agrawal1 are about the data on which the itemset operation is to be performed (col. 11, ln. 37-50; Table 1). Data characteristics are based on information such as the number of records, the number of transactions, and the number of frequent items. Data characteristics are not analogous to how busy a computer system currently

is. As a non-limiting example, a computer system that is not very busy may be able to efficiently perform a technique with high resource demands on a large dataset. On the other hand, a computer system that is very busy may not be able to efficiently perform the technique with high resource demands on the large dataset. Information about the dataset alone is not sufficient to make a selection of an occurrence counting technique that is based on conditions of the computer system such as how busy the computer system currently is.

**Further, the data characteristics of Agrawal1 do not reflect how busy a computer system is "before a frequent itemset operation is performed."** At best, information about an operation to be handled on a computer system is related to a future condition of the computer system as the operation is performed. If the operation is on a large dataset, then the computer system is more likely to be busy processing the operation **in the future**. If the operation is not on a large dataset, then the computer system is less likely to be busy processing the operation **in the future**. Before the operation is performed, the data characteristics alone provide no information about the current busyness of the computer system. Making a determination based on the data characteristics alone may cause a technique with high resource demands to be performed even though the computer system is very busy before the technique is selected.

Agrawal1 also mentions that a disadvantage to the GatherCount approach is that the system may run out of memory as GatherCount is being performed (col. 11, ln. 17-22). Specifically, Agrawal1 states: "If enough memory is not available, $C_2$ needs to be partitioned and the process has to be repeated for each partition" (col. 11, ln. 17-22). **Running out of memory while performing an approach is not the same as selecting an approach based on how busy the computer system currently is, before the frequent itemset operation is performed.** Agrawal1 does not mention a system that could avoid selecting the GatherCount

approach based on how much memory is available. Instead, Agrawal1 slavishly executes GatherCount even though there is not enough memory available to perform GatherCount in memory. Unlike Agrawal1, Claim 37 comprises selecting an occurrence counting technique "based on conditions existing before a frequent itemset operation is performed in a computing environment in which the frequent itemset operation is to be performed, wherein the conditions include how busy a computer system in which the frequent itemset operation is to be performed currently is."

Finally, **Agrawal1 does not involve "dynamically selecting" an occurrence counting technique. Instead, Agrawal1 mentions a technique for analyzing results of an experiment to find the best static approach to use (col. 14, ln. 36-46)**. Table 1 is provided for the purpose of "cost analysis of different approaches" (Table 1) to be analyzed. Based on Table 1, a comparison of different approaches that resulted from "experiments" is shown in Table 2 (col. 14, ln. 4). Agrawal1 concludes the experiment by selecting the Vertical approach as the best approach (col. 14, ln. 36-46). Nothing in the art of record mentions "dynamically selecting" an occurrence counting technique in the manner recited in Claim 37.

Agrawal2 is merely cumulative with Agrawal1 and does not fill the gaps left by Agrawal1. Agrawal2 mentions testing execution plans that use different data structures, such as indexes and materialized views (col. 2, ln. 51-col.3, ln. 11). The test results are used to estimate the cost of using the different data structures. Based on the estimated costs of using an index, a determination may be made as to "which set of indexes and materialized views provides the most benefit." **Information about past results of using logical arrangements of data has nothing to do with dynamically selecting an occurrence counting technique "based on conditions existing before a frequent itemset operation is performed in a computing environment in**

**which the frequent itemset operation is to be performed." Further, the test results do not show conditions that include "how busy a computer system in which the frequent itemset operation is to be performed currently is."** At best, the results are indicative of how busy a computer system previously was. For this reason, Agrawal2 is merely cumulative with Agrawal1.

The Final Office Action erroneously cites Agrawal2 as showing conditions that include "how busy a computer system currently is." Because of this clear factual error, the Final Office Action fails to support its position that the features of Claim 37 are obvious.

Further, Claim 37 includes substantial features discussed above that are not shown, discussed, or even hinted at in the art of record, and these features would not have been obvious to a person of ordinary skill in light of the references. For at least those reasons stated, Claim 37 is patentable over the art of record, taken individually or combined. Accordingly, Applicant respectfully requests withdrawal of the rejection under 35 U.S.C. § 103(a) of Claim 37, which is believed to be in condition for allowance.

CLAIMS 39-40, 47, AND 49-50

Claims 39 and 40 depend from Claim 37 and are patentable over the art of record for at least those reasons discussed with respect to Claim 37. Claims 47 and 49-50 are computer-readable storage medium claims that include the non-obvious subject matter recited in Claim 37. Thus, Claims 47 and 49-50 are patentable over the art of record for at least those reasons provided with respect to Claim 37. Applicant respectfully requests withdrawal of the rejection under 35 U.S.C. § 103(a) of Claims 39-40, 47, and 49-50, which are believed to be in condition for allowance.

INDEPENDENT CLAIM 38

As discussed in the interview, the amendments to Claim 38 do not affect the allowable subject matter in the claim. Applicant respectfully requests that the amendments be entered to efficiently advance this application towards a Notice of Allowance.

Claim 38 includes each of the features recited and discussed with respect to Claim 37, and Claim 38 is patentable over the art of record for at least those reasons provided above with respect to Claim 37.

Additionally, Claim 38 features other patentable subject matter that is discussed in more detail below.

The method of Claim 38 features different occurrence counting techniques that are dynamically selected for an N-item phase at different times. As a non-limiting example, at time $t_1$ a first technique is selected for a 3-item phase when a computer system is very busy, and at time $t_2$ a second technique is selected for the 3-item phase when the computer system is not very busy. As stated in Applicants' specification, "while a first technique may perform vastly superior to a second technique under certain conditions, the second technique may perform vastly superior to the first technique under other conditions" (par. [0184]).

Specifically, Claim 38 features "dynamically selecting" "a first occurrence counting technique" for "a phase during which combinations having N items are processed." A determination is made as to whether "candidate combinations" of items for the phase "satisfy . . . frequency criteria." Then, Claim 38 features "dynamically selecting a second occurrence counting technique," "different from" the first occurrence counting technique, "in the phase of a subsequent frequent itemset operation during which combinations having N items are processed."

Agrawal1 mentions computing the cost of an approach either after the approach is fully performed or after each pass of the approach (col. 11, ln. 40-44) as part of an experiment (col. 14, ln. 1-13). Agrawal1 does not mention "dynamically selecting" an occurrence counting technique for an N-item "phase." Although costs of each pass may be computed and analyzed, Agrawal1 merely mentions comparing different approaches and statically choosing the best approach to use each time (col. 14, ln. 36-46). For example, Agrawal1 concludes that the Vertical approach has the best performance (col. 14, ln. 36-46). Nothing in the art of record suggests "dynamically selecting" an approach for an N-item "phase."

Further, Agrawal1 does not mention different occurrence counting techniques that are selected for an N-item phase at different times. Agrawal1 mentions finding the best approach by calculating the cost of various approaches. Table 1 analyzes the costs of different approaches by presuming that costs depend only on data characteristics. For example, Table 1 shows a cost of joining two relations of size n and m to get a result of size r that has a specified cost. The specified cost in the table does not depend on when the join is performed, and Table 1 has nothing to do with conditions of a computer system at varying times. Unlike Agrawal1, Claim 38 features selecting occurrence counting techniques based on conditions that include how busy a computer system currently is, wherein different techniques are selected for an N-item phase at different times.

The art of record does not show at least the recited features, and Claim 38 is in condition for allowance.

CONCLUSION

For the reasons set forth above, it is respectfully submitted that all of the pending claims are now in condition for allowance. Therefore, the issuance of a formal Notice of Allowance is believed next in order, and that action is most earnestly solicited.

The Examiner is respectfully requested to contact the undersigned by telephone if it is believed that such contact would further the examination of the present application.

Please charge any shortages or credit any overages to Deposit Account No. 50-1302.

Respectfully submitted,

Hickman Palermo Truong & Becker LLP

Dated: October 2, 2009

/EricLSutton#61173/
Eric L. Sutton
Reg. No. 61,173

2055 Gateway Place, Suite 550
San Jose, California 95110-1089
Telephone No.: (408) 414-1080
Facsimile No.: (408) 414-1076